

International Conference on Modeling, Optimization and Computing (ICMOC)

An Interactive system for Sensory and Gustatory Impaired People Based On Hand Gesture Recognition

Elakkiya R^a, Selvamani K^b, Kanimozhi S^c,
Velumadhava Rao^d, Senthilkumar J^{e,a*b*}

^{a,d}Student, DCSE, Anna University, Chennai-25, India

^bAsst.Professor, DCSE, Anna University, Chennai-25, India

^cSystem Analyst, Ramco Systems, Chennai-113, India

^eResearch Scholar, DCSE, Anna University, Chennai-25, India

Abstract

Sign Language is an interactive system for sensory and gustatory impaired people for communicating information among them. Hand Gesture Recognition (HGR) is a system for sign language detection which uses computers for enabling communication with sensory impaired people. In this research work, an efficient and effective technique is introduced for the identification of the number of fingers opened or closed in a gesture representing an alphabet of the American Sign Language (ASL). Finger Detection is accomplished with the use of Contour Hull and Convexity Defects. Contour Hull helps to form the boundary of the hand gestures whereas the Convexity Defects helps to find tip of the fingers captured by the web camera. This proposed system applies image processing techniques to store the pre-processed data, Machine learning technique is used in supervised clustering and computer vision techniques is used to recognize the hand gestures. Hence this intelligent system does not require the hand to be perfectly aligned to the camera for detecting the signs. In addition to this, to detect the signs, this system does not need any colored markers or gloves to be wearied on the hands.

© 2012 Published by Elsevier Ltd. Selection and/or peer-review under responsibility of Noorul Islam Centre for Higher Education Open access under [CC BY-NC-ND license](#).

Keywords: Skin Detection; Finger Detection; Machine Learning; Computer Vision Techniques; Sign Language Recognition;

* Corresponding author. Tel.: +91-9940658410, +91-9940098167; fax: 044-22358800.

E-mail address: elakkiyaceg@gmail.com, smani@cs.annauniv.edu, kanimozhis@gmail.com, reachvelurao@gmail.com, senthilkumar@cs.annauniv.edu

1. Introduction

Computer Recognition of Sign Language is an important issue for enabling communication with hearing impaired people. An efficient and fast algorithm for identification of the number of fingers in a gesture representing an alphabet of the American Sign Language is proposed in this work. Finger Detection is achieved by the method of Boundary Tracing and Finger Tip Detection. This proposed system does not require the hand to be perfectly aligned to the web camera or use any special markers and input gloves on the hand. Moreover, Computer Vision techniques are adapted for the recognized hand gestures which are assigned to the mouse control to perform cursor movements and make the interaction with the systems more powerful.

In general, Deaf community uses sign language as the main communication tool for communicating their information. Sign language gives more priority for visual communication language. Users are using the identification namely orientation, shape and movement of the hands, arms, body, and facial expressions to reveal their expression to interact with the sensory impaired people. The current ASL system is used for all communication between the impaired people. But the way that the communication tool developed was complicate and restricts the communication with other normal people by reducing the understanding level of the public. To overcome the limitations of the current system communication, the existing system needs sign language interpretation (specially used by the hearing impaired) that makes the language easily understood by the people.

This proposed technique is effective to capture the hand position, extract the shape of the hand and to classify the signs captured from the users by the web camera. Moreover the captured images are used to find the location of the hands in each frame, then the Haar Classifier is generated to verify whether the frame captured contain hand gesture or other object trained in the supervised learning. Hand shape is extracted by using skin detection and noise removal methods followed by thresholding and normalization. In addition, the binary image of the hands is classified based on the collected images of hand gestures used in training data. Classification algorithm used to classify the detected gestures is K Nearest Neighborhood algorithm. The value of the accuracy can vary depending on the consistency of the resulting training data and noise.

A hand posture, gesture modelling and recognition system is used to determine the shape of the detected hand based on the coordinate model computed by the hand detection and tracking system which is proposed in this work. Hand gesture trajectory recognition is introduced to interact with the sensory impaired people. Dynamic hand gesture recognition is made along with the skin detection of different racial people. Colour tracking is performed to differentiate the skin colour from a group of Red Green Blue (RGB) colours. Hand gestures are recognized using multiple frames to detect its trajectory direction. All the frames are assigned with the instructions in the *Knowledge Base* and stored in the *Gesture Base* as the pre-processed data. These datasets are used to perform necessary interfaces with the computers based on the hand motions captured by the web camera.

In this proposed work, the recognized hand gestures are used to perform different operations in two applications namely Sign Language Recognition (SLR) and Mouse Move over Control (MML). But in this paper, we are concentrated only on Sign Language Recognition. The proposed system architecture is shown in Figure 1. This architecture consists of three major components namely image processing components, supervised learning and unsupervised learning to detect the hand gestures effectively. More over to extract hand features, the necessary hand contour is evaluated and to reduce noise in the captured frames the morphological operations (erosion and dilation) are performed. These two processes are involved in the image processing components at the first stage. Positive and negative samples are trained by the supervised learning technique with Haar-Classifer. With the trained gestures, unsupervised learning will be adapted to classify the hand gestures using K Nearest Neighbourhood (KNN) algorithm. Also, the Region of Interest (ROI) is estimated to perform the necessary mouse operations for various interactions.

2. Related Work

There are many works carried out in the past by many researchers pertaining to Artificial Neural Networks in hand gesture based recognition. Some of them are cited in this literature survey. Among them, Annamária R. et.al introduced a method of combining artificial neural networks and fuzzy logic to reduce the training time which results in better accuracy. In their work, the trained data are clustered using the neural nets and centers is introduced. Ruiduo Yang et.al considered two crucial problems in continuous sign language recognition from unaided video sequences.

V. López-Ludeña et.al presented a rule-based translation method and a statistical translator. Their evaluation was carried out in the Local Traffic Office in the city of Toledo (Spain) involving real government employees and deaf people. Ehsan ul haqI et al. combined existing techniques of skin color based ROI segmentation and Viola-Jones Haar-like feature based object detection, to optimize hand gesture recognition for mouse operation.

Another technique was introduced by Bo Peng, and Gang Qiasn called video-based framework for view-invariant and full-body gesture spotting. In addition to this, they also presented extracting view-invariant pose features using multi-linear analysis from visual hull data, Hidden Markov Models (HMMs) are trained for gesture recognition by using these pose features as observations. Hand posture and gesture modeling and recognition system was introduced by Annamária R. et al, which is used as an interface to make possible communication with smart environment by simple hand gestures.

3. Sign Language Recognition

The proposed system architecture for Sign Language Recognition is shown in figure 1. This architecture consists of various components namely Skin Detection, Finger Detection, Morphing Operation, Haar-Classifer, Machine Learning and Computer Vision Techniques. Alphabets representing the American Standard Language are collected from different racial people to detect the skin tone. This system is able to differentiate the skin color using histogram techniques. From the collected images the contour is formed with the help of Haar-classifier. Boundary tracing is done for the exact shape of the hand gestures to form the positive and negative samples.

This proposed system is divided into three modules namely Image Processing Module, Training Module and Exploitation Module. First all the signs are collected and stored in the database for training. Second, positive and negative samples are differentiated and trained using supervised clustering algorithm. Finally with the help of computer vision techniques, hand gestures are recognized and the corresponding signs are identified. The detailed explanations of each module are listed below.

3.1. Skin Detection

Skin detection module is used to separate the skin tone from the group of RGB colours in the frames. Skin color detection is implemented by enhancing the peer model with HSV color space. The enhancement of this peer model with normalization is designed to get the average skin colour of the people from various racial majorities, and also under different lighting conditions gives better results.

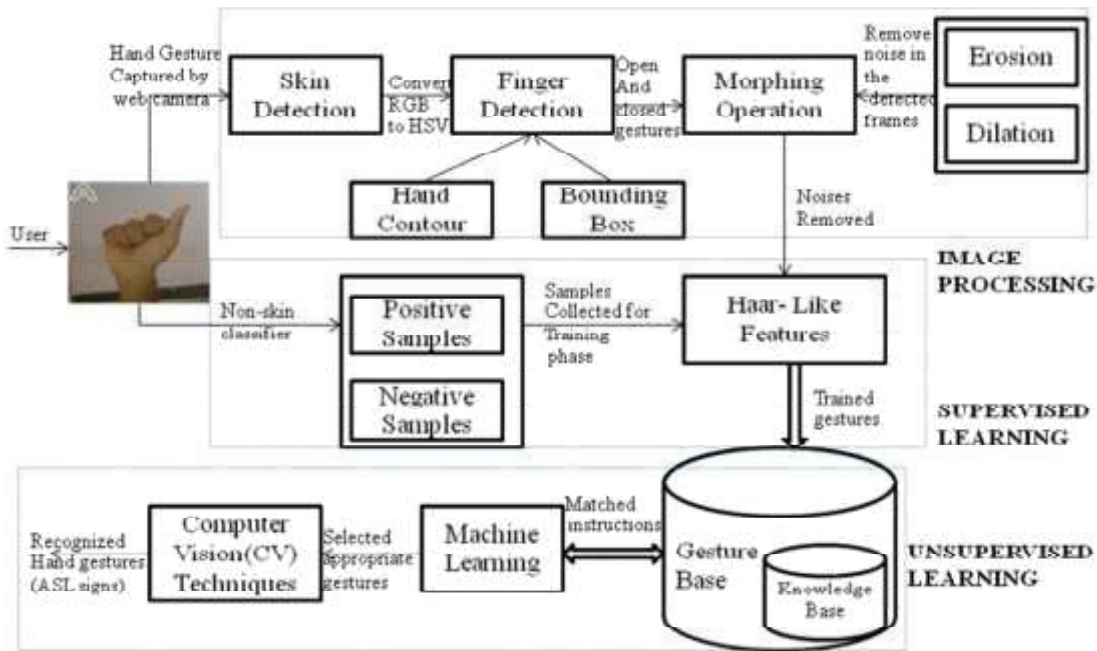


Fig.1 Architecture for Sign Language Recognition

This enhanced peer model is represented with the following values and conditions. The values of RGB is set to the condition (Eqn.1 & 2) such that

$$R > 95, G > 40 \text{ and } B > 20 \quad (1)$$

Also,

$$\text{Max}(R, G, B) - \text{Min}(R, G, B) > 15 \quad (2)$$

$$\text{Abs}(R - G) > 15 \text{ with } R > G, R > B \text{ and } G > B.$$

Where R=red, G=green and B=blue.

Finally, the HSV values (Eqn. 3) are set to

$$H < 19, S > 48 \text{ and } V > 35 \quad (3)$$

Where H=hue, S=saturation and V=Value.

3.2. Finger Detection

Fingers in the hand are detected using the edge detection algorithm. In finger detection, first the boundaries are traced for the hand gestures shown in front of the web camera. Second hand contours are formed around the recognized hand gestures. Boundary tracing helps to form the hand contour. Finally, this contour along with the convexity defect finds the number of opened and closed fingers in the hand gesture which is shown in front of the camera. The various stages of finger detection module are depicted in the figure 2.

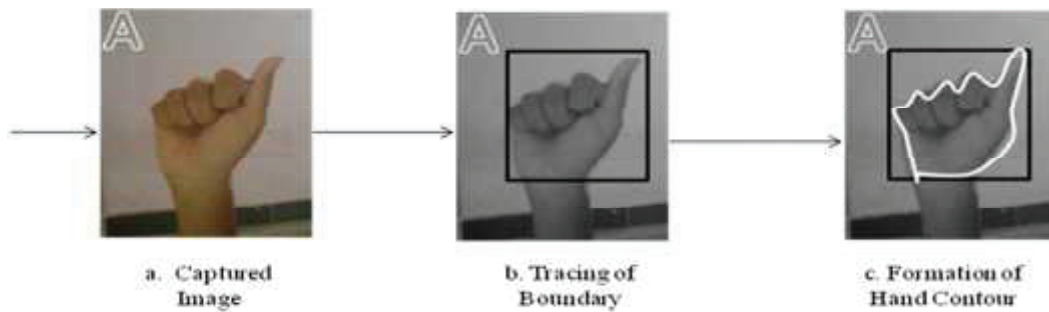


Fig.2 Finger Detection Process

3.3. Haar-Classifier

The Haar-Classifier module is used to detect other objects (i.e. negative samples apart from hand gestures) in the frames that are currently detected by the web camera. This method detects the objects and it is used to build a boosted rejection cascade, which will discard the negative data. Hence a negative training decision is obtained in order to determine the positive data. Haar Classifier is also used in supervised learning to detect the hand gestures and other objects at the time of training. Haar-classifier objects can be of any types such as rectangle, square with respect to edges or decision box. These kinds of classifier are trained for negative data during the training phase.

For calculating the value of the rectangle, Integral Image is used. It is the central representation of an image that consists of sum of the value in gray-scale pixels. Once the value of the integral image is obtained, it has to find a cascade classifier which is a chain stage classifier. The chain stage classifier is used to detect the sub window, within which the image contained the desired object. Hand gestures with the positive data are trained by means of clustering. These data are clustered with the help of k-means clustering algorithm. To reduce the training time the data can be clustered during the data collection stages by using Fuzzy Neural Networks (FNN) concepts. Once the data is collected with the centroids of clustering, it is easier to analyze the positive trained samples in the database when the hand gesture is identified.

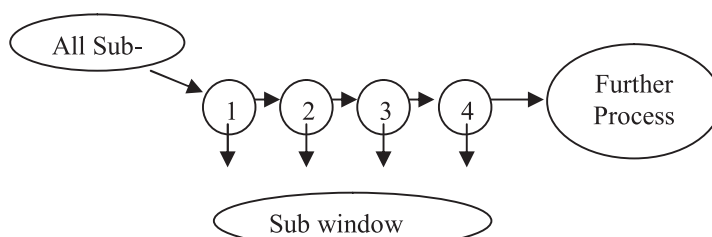


Fig. 3 Tree Cascade Classifier

3.4. K-Nearest Neighbour Classifier

The K-nearest neighbor algorithm is a technique that classifies the detected objects based on learning data which are located nearest to the object. Learning data is projected into many dimensions of the space, where

each dimension represent features of the data. The space is divided into different sections based on the classification of trained data. A point in the space which is most nearest to the clustered data will matches with the centroids to find the nearest value of the neighbors. Nearest or distant neighbors are calculated using the Euclidean distance algorithm.

In the training phase, this algorithm performs only the storage and feature vectors classification of the learning data. In exploitation phase, the same features for the test data (in which the classification is not known) is calculated as the distance from the new vector, the data vectors and the number of the closest points are considered in this phase. A new point classified and predicted to include the classification of most trained data points as shown in figure 3. The accuracy of K-NN algorithm is much influenced with the presence or absence of the irrelevant features that irrelevant or if the weight of the feature is not equivalent with relevance to the classification. When the amount of data approaches the level of infinity, this algorithm also ensures the error rate is reduced much when compared with Bayesian network error rate.

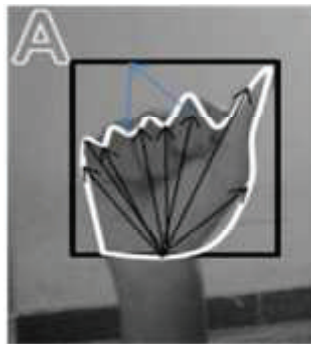


Fig.4 K-Nearest Neighbors

4. Implementation Setup

The experimental setup for analyzing the efficient interaction of hand gestures is described as follows. The web camera can be mounted in our gestures to project the information onto the surface walls. Web camera has to be placed to recognize the skin colour and trajectory direction of hand gestures which depends upon the resolution of the camera. Depends upon the resolution of the camera, the accuracy of the motion detection will be increased and the camera is placed within the distance according to the range of the pixels of hand gestures. The algorithms are running in a PC and can be implemented using *Microsoft Visual studio 2010*. Once the image is obtained from the camera, it is processed to perform skin color detection and proceed to filter the hand gestures for necessary operations. The desired object is obtained and with that the data classification is classified and stored in the gesture base with the necessary instructions.

Throughout the process, the following parameters have to be considered. The input frames can be 8bit/channel with dynamic hand gestures at the resolution of 640 X 480 pixels. The images trained with Haar-Classifiers are single channelled and considered with same parameters of light intensity. Once the skin pixels are detected, only the pixels having brightness (i.e. Value in the HSV space) has to be retained. Others which has the value less than the thresholding parameter has to be discarded. The experimental setup is viewed in two phases to implement the sign gesture recognition. i.e. A. Supervised Learning B. Unsupervised learning or machine learning.

5. Conclusion

After carrying out the testing methodologies for the system "Sign Language detection" which detects the hand gestures with the maximum distance of 90 cm. The optimal results are obtained by detecting the hand through the web camera in the exact position. The modified approach that uses HSV color model detects the skin colour accurately with the RGB colour model. Moreover, the noises are completely removed which in turn produces effective results in accordance with consistent training data. When KNN algorithm is used for classification which results in good accuracy in detecting the hand gestures and hence it is proved that the detection ratio is higher in comparison with other systems.

References

- [1] Annamária R. Várkonyi-Kóczy, and Balázs Tusor . Improved back-propagation algorithm for neural network training., *IEEE Trans. on Intelligent Signal Processing* ; Oct 2011, vol.39, No.5, pp. 504 – 511.
- [2] Ruiduo Yang, Sudeep Sarkar, Senior Member, IEEE, and Barbara Loeding. Handling Movement Epenthesis and Hand Segmentation Ambiguities in Continuous Sign Language Recognition Using Nested Dynamic Programming. *IEEE Trans. On Pattern Analysis and Machine Intelligence*; Mar 2011, Vol.32, No.3, pp. 462-477.
- [3] Gaurav Pradhan and Balakrishnan Prabhakaran. Hand-Gesture Computing for the Hearing and Speech Impaired. *Proc. of IEEE 8th Conf. on Machine Intelligence*; Aug 2011.
- [4] V. López-Ludeña, R. San-Segundo, R. Martín, D. Sánchez and A. García. Evaluating a Speech Communication System for Deaf People. *IEEE Trans. on Latin America* May 2011; Vol.9, no.4, pp. 565-570.
- [5] Ehsan ul haqI, Syed Jahanzeb, Hussain PirzadeI, Mirza Waqar Bail, and Hyunchul Shin. New Hand Gesture Recognition Method for Mouse Operations. *Proc. of IEEE 54th International Midwest Symposium on Circuits and Systems (MWSCAS)*; Sep 2011, pp.1-4.
- [6] Chen-Chiung Hsieh, Dung-Hua Liou, and David Lee. A Real Time Hand Gesture Recognition System Using Motion History Image. *Proc. IEEE Int'l conf. Signal Processing Systems (ICSPS)*; 2010; pp.394-398.
- [7] M.K. Bhuyan, Debanga Raj Neog and Mithun Kumar Kar. Hand Pose Recognition Using Geometric Features ,*proc. IEEE Conf. on fuzzy systems*; Mar 2011, Vol.23, No.5.
- [8] Bo Peng, and Gang Qian. Online gesture spotting from visual hull data. *IEEE Trans. Pattern Analysis And Machine Intelligence*, June 2011; Vol. 33, No. 6.
- [9] Kabeer Manchanda, and Benny Bing. Advanced Mouse Pointer Control Using Trajectory-Based Gesture Recognition. *Proc. IEEE southeast conf on motion control*; Apr. 2010, pp 412-415.
- [10] Tan Wenjun, Wu Chengdong, Zhao Shuying and Jiang Li. Dynamic Hand Gesture Recognition Using Motion Trajectories and Key Frames. *IEEE Trans. Image and Vision Computing*; oct 2010, Vol.21, No.8.
- [11] Thomas Hahn. Future Human Computer interaction with special focus on input and output techniques. *IEEE Trans. On Systems, man and cybernetics*; MARCH 2010 pp.1-13.
- [12] Fariborz Mahmoudi, Mehdi Parviz. Visual Hand Tracking Algorithms. *Proc. IEEE Conf. of the Geometric Modelling and Imaging— New Trends (GMAI'06)*; JULY 2006, pp. 323-327.